Published in final edited form as:

Biomol NMR Assign. 2015 October ; 9(2): 431–434. doi:10.1007/s12104-015-9624-0.

^1H , ^{15}N and ^{13}C backbone resonance assignments of the N-terminal, tandem KH domains of human hnRNP E1

Yang Li¹ and Mirko Hennig^{2,*}

¹Department of Biochemistry and Molecular Biology, Medical University of South Carolina, Charleston, South Carolina 29425, USA

²Department of Nutrition, University of North Carolina at Chapel Hill/Nutrition Research Institute, Kannapolis, North Carolina 28081, USA

Abstract

Heterogeneous nuclear ribonucleoproteins (hnRNPs) can be divided into subgroups based on their RNA-binding characteristics. One subgroup in mammalian cells are the Poly(C)-binding proteins (PCBPs) comprised of hnRNP K/J and hnRNP E1-4 (the latter also known as PCBP 1-4 or αCP [α -complex proteins] 1-4). Each subgroup member has three K homology (KH) nucleic acid-binding domains. Individual KH domains bind short single-stranded (ss), poly-pyrimidine-rich nucleic acid sequences with rather weak affinity. In this study, we report the ^1H , ^{13}C and ^{15}N backbone resonance assignments of the first and second KH domains of hnRNP E1, which plays a pivotal role in posttranscriptional and translational regulation of RNA targets. Our NMR assignments lay the foundation for a detailed investigation of the dynamic cooperation of the tandem KH1 and 2 domains to bind nucleic acids.

Keywords

Heterogeneous nuclear ribonucleoprotein E1 (hnRNP E1); Poly(C)-binding protein 1 (PCBP1); K homology (KH) domain; nucleic acid binding

Biological context

hnRNP proteins play important roles in the regulation of gene expression by interacting with both, RNA and DNA targets (Matunis, Michael et al. 1992), Leffers, Dejgaard et al. (1995). For example, hnRNP K binds to the single stranded CT element of human c-myc gene and activates its transcription (Tomonaga and Levens 1996). On the other hand, hnRNP E1 mediates the translational repression of Disabled-2(Dab2) and interleukin-like EMT inducer (ILEI) through interactions with an RNA element in the 3'-UTR of those mRNA transcripts, termed BAT (TGF β -activated translation) element (Hussey, Chaudhury et al. 2011).

Each member of the hnRNP K/J and E1-4 subgroup contains three K homology (KH) domains which consist of ca. 70 amino acids. These KH domains have the classical type I

* Address correspondence to: Dr. Mirko Hennig, University of North Carolina at Chapel Hill, Nutrition Research Institute, Room 3151, 500 Laureate Way, Kannapolis, North Carolina 28081, USA, tel.: (704) 250 5060, mirko_hennig@unc.edu.

KH fold found in eukaryotes and feature a $\beta 1$ - $\alpha 1$ - $\alpha 2$ - $\beta 2$ - $\beta 3$ - $\alpha 3$ topology (Grishin 2001). As gene expression regulators, hnRNP K/J and hnRNP Es preferentially bind to C-rich sequences of ss-target nucleic acids through their modular KH domains. Typically, only four nucleobases are accommodated by this KH-domain binding cleft. The oligonucleotide binding groove is comprised of helices $\alpha 1$ and $\alpha 2$ on one side and strands $\beta 2$ and $\beta 3$ on the other (Yoga, Traore et al. 2012). The highly conserved GXXG motif connecting helices $\alpha 1$ and $\alpha 2$ and the second loop linking strands 2 and 3 is also crucial to binding.

The repetitive KH nucleic acid-binding modules of hnRNP E1 and its functionally non-redundant homologue, hnRNP E2 (Du, Fenn et al. 2008), are spatially distributed such that the two N-terminal KH domains form one structural unit with an interface comprised of strands $\beta 1$ and helix $\alpha 3$ of each individual domain. On the other hand, the third KH domain at the C-terminus is separated from the second KH domain by a 110 residue long linker (Makeyev and Liebhaber 2002). The preformed platform involving the first two KH domains and the flexible third KH domain may either act independently or cooperatively to increase nucleic acid binding specificity and/or affinity. To better understand how hnRNP E1 interacts with nucleic acids, we report the ^1H -, ^{13}C - and ^{15}N -NMR backbone resonance assignments of the tandem KH1 and KH2 domains of hnRNP E1. This will serve as a basis for further investigation of KH1-KH2 dynamics and its interactions with nucleic acids.

Methods and experiments

The 15kDa, tandem KH1-KH2 domains of hnRNP E1 were cloned into the expression vector pET-23a(+) carrying a C-terminal HisTag sequence. Stabilizing mutations of C54S, C109A, C118A, C158S and C163S were introduced. Uniformly $^{13}\text{C}/^{15}\text{N}$ - and ^{15}N -labeled KH1-KH2 NMR samples were produced in *Escherichia coli* SOLUBL21(DE3)-RIPL cells. The bacterial cultures were grown in M9 minimal medium with ^{13}C -glucose and/or $^{15}\text{NH}_4\text{Cl}$ as the sole sources for carbon and nitrogen, respectively. Protein expression was induced with 1mM IPTG overnight at 18 °C. Subsequently, samples were purified using HisTrapFF affinity chromatography (GE Healthcare) followed by size exclusion chromatography employing a Superdex 75 26/60 column. Samples were concentrated to ca. 0.3 mM in buffer containing 50mM sodium acetate (pH 5.4), 100 mM EDTA and 50 μM sodium azide in 90%/10% $\text{H}_2\text{O}/\text{D}_2\text{O}$ or 99% D_2O for NMR spectra acquisition.

NMR data for the backbone resonance assignment of KH1-KH2 were collected on uniformly $^{13}\text{C}/^{15}\text{N}$ -, or ^{15}N -labeled samples. Spectra were recorded at 298K on a Bruker Avance800 equipped with TCI-cryoprobe. Sequential backbone resonance assignments for ^1HN , ^{15}N , $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$ and $^{13}\text{C}'$ were obtained using 2D ^1H - ^{15}N SOFAST-HMQC and ^1H - ^{13}C HSQC as well as 3D HCCH-TOCSY, HNCACB, CBCA(CO)NH, HNCA and HN(CO)CA experiments. Subsequently, assignments were confirmed using 3D ^{15}N -resolved HSQC-NOESY-HSQC and NOESY-HSQC experiments. 3D HBHA(CO)NH and HNHA were employed for $^1\text{H}\alpha$ and $^1\text{H}\beta$ chemical shifts assignments. The NMR data were processed using nmrPipe, and analyzed with CcpNMR Analysis software (Vranken, Boucher et al. 2005).

Assignments and data deposition

An assigned 2D ^1H - ^{15}N SOFAST (Band-Selective Optimized Flip Angle Short Transient) HMQC spectrum of KH1-KH2 is shown in Fig. 1. In total, 95% backbone resonance were assigned, including 97% of the H^{N} and N amide protons and nitrogens, 93% of $\text{C}\alpha$ -, 94% $\text{C}\beta$ -, and 93% C' -carbons. In addition, 95% of $\text{H}\alpha$ - and 92% of $\text{H}\beta$ -proton assignments were obtained. Assignment of E34, located in a flexible loop, is missing due to unfavorable exchange dynamics. While ^1H chemical shifts were externally referenced to DSS, heteronuclear ^{13}C and ^{15}N chemical shifts were referenced indirectly according to their $\text{X}/^1\text{H}$ ratio in DSS. Chemical shift assignments have been deposited in the BioMagResBank database (<http://www.bmrb.wisc.edu>) under accession number 25494.

KH1-KH2 secondary structure elements were predicted from chemical shifts using TALOS+ as shown in Fig. 2 (Shen, Delaglio et al. 2009). The resulting secondary structure elements confirm the classical KH domain type I secondary structure composition. The first KH domain includes three α -helices ($\alpha 1$: residues 23–29, $\alpha 2$: residues 33–43, $\alpha 3$: residues 65–85) and three β -strands ($\beta 1$: residues 14–21, $\beta 2$: residues 46–49, $\beta 3$: residues 57–62). The second KH domain is composed of three α -helices ($\alpha 1$: residues 106–112, $\alpha 2$: residues 117–126, $\alpha 3$: residues 152–169) and three β -strands ($\beta 1$: residues 98–103, $\beta 2$: residues 130–134, $\beta 3$: residues 144–150).

Acknowledgments

This work was supported by the South Carolina Clinical & Translational Research (SCTR) Institute (UL1TR000062). We thank Dr. Daniella Ishimaru (MUSC) for KH1-KH2 NMR sample preparation and acknowledge Dr. Philip H. Howe (MUSC) for the generous gift of the parent GST-hnRNP E1 construct and the support of the Hollings Marine Laboratory NMR facility for this work.

References

- Du Z, Fenn S, Tjhen R, James TL. Structure of a construct of a human poly(C)-binding protein containing the first and second KH domains reveals insights into its regulatory mechanisms. *J Biol Chem.* 2008; 283(42):28757–28766. [PubMed: 18701464]
- Grishin NV. KH domain: one motif, two folds. *Nucleic Acids Res.* 2001; 29(3):638–643. [PubMed: 11160884]
- Hussey GS, Chaudhury A, Dawson AE, Lindner DJ, Knudsen CR, Wilce MC, Merrick WC, Howe PH. Identification of an mRNP complex regulating tumorigenesis at the translational elongation step. *Mol Cell.* 2011; 41(4):419–431. [PubMed: 21329880]
- Leffers H, Dejgaard K, Celis JE. Characterisation of Two Major Cellular Poly(rC)-Binding Human Proteins, Each Containing Three K-homologous (KH) Domains. *Eur J Biochem.* 1995; 230(2):447–453. [PubMed: 7607214]
- Makeyev AV, Liebhaber SA. The poly(C)-binding proteins: a multiplicity of functions and a search for mechanisms. *RNA.* 2002; 8(3):265–278. [PubMed: 12003487]
- Matunis MJ, Michael WM, Dreyfuss G. Characterization and primary structure of the poly(C)-binding heterogeneous nuclear ribonucleoprotein complex K protein. *Mol Cell Biol.* 1992; 12(1):164–171. [PubMed: 1729596]
- Shen Y, Delaglio F, Cornilescu G, Bax A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR.* 2009; 44(4):213–223. [PubMed: 19548092]
- Tomonaga T, Levens D. Activating transcription from single stranded DNA. *Proc Natl Acad Sci USA.* 1996; 93(12):5830–5835. [PubMed: 8650178]

- Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins*. 2005; 59(4):687–696. [PubMed: 15815974]
- Yoga YM, Traore DA, Sidiqi M, Szeto C, Pendini NR, Barker A, Leedman PJ, Wilce JA, Wilce MC. Contribution of the first K-homology domain of poly(C)-binding protein 1 to its affinity and specificity for C-rich oligonucleotides. *Nucleic Acids Res*. 2012; 40(11):5101–5114. [PubMed: 22344691]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1.

Fully annotated 2D ^1H - ^{15}N SOFAST-HMQC spectrum of the KH1-KH2 domains of human hnRNP E1 recorded at 25°C. An enlarged view of the most crowded region of the spectrum is shown in the top-right corner. Assigned residues from backbone amide groups are indicated by residue type and number, folded peaks are labeled in italics, and non-native residues from the expression vector are marked by *.

Figure 2. Secondary structure composition of the KH1-KH2 domains derived from secondary chemical shifts. Differences of the chemical shift deviations of $C\alpha$ and $C\beta$ carbons ($\delta(C\beta) - \delta(C\alpha)$) with respect to corresponding random coil values are plotted against amino acid residue number. Secondary structures as identified by TALOS+ are indicated with boxes.